



Oscillatory profiles of positive, negative and neutral feedback stimuli during adaptive decision making



Peng Li^a, Travis E. Baker^b, Chris Warren^c, Hong Li^{a,*}

^a Brain Function and Psychological Science Research Center, Shenzhen University, Shenzhen, China

^b Department of Neurology and Neurosurgery, Montreal Neurological Institute, McGill University, Montreal, Canada

^c Department of Psychology, Leiden University, Leiden, The Netherlands

ARTICLE INFO

Article history:

Received 20 October 2015

Received in revised form 22 June 2016

Accepted 30 June 2016

Available online 1 July 2016

Keywords:

Theta

Beta-gamma

Feedback-related negativity

Neutral feedback

Reinforcement learning

ABSTRACT

The electrophysiological response to positive and negative feedback during reinforcement learning has been well documented over the past two decades, yet, little is known about the neural response to uninformative events that often follow our actions. To address this issue, we recorded the electroencephalograph (EEG) during a time-estimation task using both informative (positive and negative) and uninformative (neutral) feedback. In the time-frequency domain, uninformative feedback elicited significantly less induced beta-gamma activity than informative feedback. This result suggests that beta-gamma activity is particularly sensitive to feedback that can guide behavioral adjustments, consistent with other work. In contrast, neither theta nor delta activity were sensitive to the difference between negative and neutral feedback, though both frequencies discriminated between positive, and non-positive (neutral or negative) feedback. Interestingly, in the time domain, we observed a linear relationship in the amplitude of the feedback-related negativity (neutral > negative > positive), a component of the event-related brain potential thought to index a specific kind of reinforcement learning signal called a reward prediction error. Taken together, these results suggest that the reinforcement learning system treats neutral feedback as a special case, providing valuable information about the electrophysiological measures used to index the cognitive function of frontal midline cortex.

© 2016 Published by Elsevier B.V.

1. Introduction

Our ability to predict and evaluate the consequences of our actions is fundamental to adaptive decision making. Reinforcement learning (RL) theory holds that if an action is followed by positive feedback then that action will have a greater probability of being performed again, whereas if an action is followed by negative feedback then that action will have a lesser probability of being performed again (i.e. Thorndike's Law of Effect: Catania, 1999). But in everyday life, not all of our actions are followed by such binary consequences, but rather by uninformative events. In fact, the term *neutral operants* has long been used by RL theorists to describe responses from the environment that neither increase nor decrease the probability of a behavior being repeated (Skinner, 1938). While observations of electrophysiological activity over frontal midline cortex have motivated a wealth of experimental and theoretical analyses of RL, it remains unclear how uninformative feedback is ultimately processed during trial-and-error learning tasks.

Over the last decade, both time domain and time-frequency domain analyses of electrophysiological recordings have been increasingly used in research concerned with neural processes that differentiate

performance feedback indicating positive outcomes (e.g., monetary gain, correct feedback) from negative outcomes (e.g., monetary loss, error feedback) (Weinberg et al., 2014). In the time domain, event-related brain potential (ERP) studies have revealed a negative-going deflection in the ERP that peaks over frontal-central recording sites approximately 250 ms following feedback presentation. This feedback-locked ERP component, termed the feedback-related negativity (FRN), is typically enhanced following unexpected task-relevant events (e.g. negative feedback, errors) and is reduced or absent following positive feedback.¹ Interestingly, the few FRN studies examining neutral feedback have produced largely mixed results (Holroyd et al., 2006; Kujawa et al., 2013; Huang and Yu, 2014; Yu and Zhou, 2006). In particular, studies either report larger FRNs to neutral feedback compared to negative and positive feedback (Müller et al., 2005; Kujawa et al.,

¹ Recent evidence suggests that the difference in FRN amplitude between reward and error trials results from a positive-going deflection, the reward positivity (Rew-P), elicited by reward feedback (see Holroyd et al., 2008; Warren and Holroyd, 2012; Baker and Holroyd, 2011; Proudfit, 2015). Because the Rew-P typically occurs during the time-range of the FRN and P300, the difference-wave method is commonly used to isolate the reward positivity from other ERP components by taking the difference between the ERPs to positive and negative feedback. For the purpose of this study, we focused our analysis on condition-specific ERP effects by measuring the amplitudes of the FRN elicited by neutral, negative, and positive feedback.

* Corresponding author at: No 3688, Nanhai Road, Nanshan District, Shenzhen 518060, China.

E-mail address: lihongwrm@vip.sina.com (H. Li).

2013; Huang and Yu, 2014) or comparable FRN amplitudes between neutral and negative feedback (Holroyd et al., 2006).²

Although neutral feedback has yet been investigated in the time-frequency domain, electroencephalogram (EEG) oscillations in the theta frequency range (4–8 Hz) recorded over frontal midline areas of the scalp have been associated with outcome processing (Cavanagh and Frank, 2014; Cavanagh et al., 2012), as well as other cognitive processes related to effort, attention and motivation (for reviews Hsieh and Ranganath, 2014; Mitchell et al., 2008). Notably, more frontal midline theta power is observed following negative feedback compared to positive feedback, suggesting that this signal reflects an error-driven learning mechanism consistent with principles of reinforcement learning (Cavanagh et al., 2010). However, others have argued that frontal midline theta reflects sensitivity to important cognitive events in general rather than to errors in particular (Cavanagh et al., 2012), and signal the deployment of control (Cavanagh and Frank, 2014). Furthermore, power in the delta (1–4 Hz) and beta-gamma (20–30 Hz) frequency range have been shown to increase following positive feedback compared to negative feedback (Bernat et al., 2011; Cohen et al., 2007; Hajihosseini et al., 2012; Marco-Pallares et al., 2008). In particular, recent work has indicated that feedback-locked delta band activity that related to reward positivity appears to be specific to surprising rewards, but does not predict associated behavioral adjustments (Cavanagh, 2015). By contrast, feedback-locked beta-gamma activity is thought to reflect a salience signal and has been associated with ongoing adjustments of behavior (Hajihosseini and Holroyd, 2015) and cognitive demand (Chen et al., 2012; Gilbert and Sigman, 2007; Lee et al., 2003).

While decomposing feedback-related EEG and ERPs into spectral quantities has provided a thorough understanding of the cognitive processes underlying RL, much of the existing research on these time-frequency components often characterizes feedback-locked oscillatory activity (delta, theta, beta-gamma) as total spectral power. Unfortunately, this approach does not capture all the information available in these signals because total power within a given frequency band consists of the stimulus phase-locked part of the EEG that gives rise to the ERP, called the “evoked” power, and the non phase-locked part of the EEG that is invisible in the ERP, called the “induced” power (Tallon-Baudry and Bertrand, 1999). Importantly, current thinking holds that these activities reflect different cognitive processes, such that evoked activity reflects bottom-up neural activity, whereas induced activity is thought to reflect top-down modulation (Tallon-Baudry and Bertrand, 1999). Indeed, it was recently demonstrated that total theta power is equally sensitive to outcome valence and outcome probability, however, evoked theta power was mainly sensitive to outcome valence whereas induced theta power was mainly sensitive to outcome probability (Hajihosseini and Holroyd, 2013). The role of delta and beta-gamma band phase dynamics in feedback processing remains unknown. The difference in dominant frequencies (delta, theta, and beta-gamma) between negative and positive feedback could provide a deeper understanding of the phase dynamics (evoked vs induced) at play during trial-by-trial RL.

Furthermore, it is also important to consider the relationship between time domain and time-frequency domain measures. For example, a relationship between feedback-related delta activity and the amplitude of the P300 component has been demonstrated (e.g. Bernat et al., 2007; Cavanagh and Frank, 2014), possibly suggesting that the evoked portion of delta following feedback may contribute to feedback-related differences observed in the amplitude of the P300. However, this relationship has never been formally tested. Further, theta oscillations and the FRN have been extensively studied in parallel, decades-long literatures. Feedback-induced theta power and the FRN

occur at about the same time (200–400 ms post feedback) and share the same scalp location (over the frontal midline), suggesting a functional relationship between these two phenomena. In particular, converging evidence across multiple methodologies indicates that the anterior cingulate cortex (ACC) is the source of both frontal midline theta oscillations (Cavanagh and Frank, 2014) and the FRN (Holroyd and Yeung, 2012). Importantly, recent examinations of theta power and the FRN have provided a nuanced account about their relationship (Hajihosseini and Holroyd, 2013). Under this account, unexpected, task-relevant events elicit an ACC-dependent control process that manifests in the frequency domain as theta oscillations over frontal-central areas of the scalp (Cavanagh and Frank, 2014). In the time domain, the “evoked” portion of this theta activity that is consistent in phase across trials gives rise to the FRN (Hajihosseini and Holroyd, 2013; see also Yeung et al., 2004). Although both measures provide valuable information about cognitive function of frontal midline cortex, it has recently been argued that FRN amplitude is specifically sensitive to dopamine reinforcement learning signals whereas evoked theta power reflects the ACC response to unexpected events.

Given the relationship between theta oscillations and FRN, it is perhaps surprising that neutral feedback has not yet been investigated in the time-frequency domain. Furthermore, because of the inconsistency in FRN studies examining *neutral operants*, the functional role of the ACC in the cognitive processes that underlie reinforcement learning remains incomplete. Thus, in order to further our understanding of the cognitive processes underlying informative and non-informative feedback during RL, the electrophysiological response to the good, the bad, and the neutral needs to be further characterized. In the present study, we present an harmonious application of both ERP (i.e. FRN) and time-frequency (i.e. evoked and induced delta, theta, and beta-gamma power) approaches in an attempt to elucidate the discrete aspects of the electrophysiological dynamics between positive, negative, and neutral feedback (Holroyd et al., 2012).

2. Methods

2.1. Participants and procedure

Nineteen undergraduate students (eight males) aged 18–23 years participated in the experiment for monetary compensation. All participants had normal or corrected-to normal vision, were right-handed and had no neurological or psychological disorders. Two subjects were excluded out from the final analysis due to their poor behavioral performance. The study was approved by the local ethics committee. Participants were asked to perform a time estimation task (e.g. Miltner et al., 1997) that included neutral feedback. They were required to press the spacebar following a cue (a 1500 Hz sound that lasted 50 ms) to indicate that their estimate of 1 s had elapsed. Following their response, a feedback stimulus appeared on the screen indicating whether their estimation was correct (positive feedback, win 5 cents; a circle with a check mark), incorrect (negative feedback, 0 cents; a circle with a cross mark) or the feedback was absent (neutral feedback, either 5 or 0 cents; a circle with nothing inside). To note, participants did not know whether or not they received money following neutral feedback immediately, but would receive money for correct response (in total, 50% of the trials) at the end of the experiment. Participants were told that if their reaction time was within the time window from 900 ms to 1100 ms, they would receive positive feedback; otherwise they would get negative feedback. However, this time window narrowed by 10 ms if they responded correctly on the previous trial and widened by 10 ms if they responded incorrectly on the previous trial. For 1/3 of negative-feedback trials and 1/3 of positive-feedback trials, the appropriate feedback was randomly replaced with neutral feedback. Of the 288 trials, participants received 28% neutral feedback, 35% positive feedback and 37% negative feedback in total.

² It is interesting to note that across 5 experiments reported in Holroyd et al. (2006), the authors detailed that neutral feedback stimuli elicited larger FRNs than did the negative feedback stimuli, but was not statistically significant (see Figs. 1 and 2, Holroyd et al., 2006).

Download English Version:

<https://daneshyari.com/en/article/930648>

Download Persian Version:

<https://daneshyari.com/article/930648>

[Daneshyari.com](https://daneshyari.com)